Probability and Stochastic Processes

A Friendly Introduction for Electrical and Computer Engineers

SECOND EDITION

Roy D. Yates David J. Goodman

Definitions, Theorems, Proofs, Examples, Quizzes, Problems, Solutions

Chapter 6

Section 6.1

Expected Values of Sums

Theorem 6.1

For any set of random variables $X_1, \ldots, X_n$, the expected value of $W_n = X_1 + \cdots + X_n$ is

$$E[W_n] = E[X_1] + E[X_2] + \cdots + E[X_n].$$

Proof: Theorem 6.1

We prove this theorem by induction on n. In Theorem 4.14, we proved $E[W_2] = E[X_1] + E[X_2]$. Now we assume $E[W_{n-1}] = E[X_1] + \cdots + E[X_{n-1}]$. Notice that $W_n = W_{n-1} + X_n$. Since W_n is a sum of the two random variables W_{n-1} and X_n, we know that $E[W_n] = E[W_{n-1}] + E[X_n] = E[X_1] + \cdots + E[X_{n-1}] + E[X_n]$.

Theorem 6.2

The variance of $W_n = X_1 + \cdots + X_n$ is

$$\text{Var}[W_n] = \sum_{i=1}^{n} \text{Var}[X_i] + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{Cov}\left[X_i, X_j\right].$$

Proof: Theorem 6.2

$$Var[W_n] = E\left[\left(\sum_{i=1}^{n}(X_i - \mu_i)\right)^2\right] = E\left[\sum_{i=1}^{n}(X_i - \mu_i)\sum_{j=1}^{n}(X_j - \mu_j)\right]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} E\left[(X_i - \mu_i)(X_j - \mu_j)\right]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\begin{bmatrix} E\left[(X_1-\mu_1)^2\right] & E\left[(X_1-\mu_1)(X_2-\mu_2)\right] & \cdots & E\left[(X_1-\mu_1)(X_n-\mu_n)\right] \\ E\left[(X_2-\mu_2)(X_1-\mu_1)\right] & E\left[(X_2-\mu_2)^2\right] & \cdots & E\left[(X_2-\mu_2)(X_n-\mu_n)\right] \\ \vdots & \vdots & \ddots & \\ E\left[(X_n-\mu_n)(X_1-\mu_1)\right] & E\left[(X_n-\mu_n)(X_2-\mu_2)\right] & \cdots & E\left[(X_2-\mu_2)^2\right] \end{bmatrix}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov\left[X_i, X_j\right]$$

$$= \sum_{i=1}^{n} Var\left[X_i\right] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Cov\left[X_i, X_j\right]$$

Theorem 6.3

When $X_1, \ldots, X_n$ are uncorrelated,

$$\text{Var}[W_n] = \text{Var}[X_1] + \cdots + \text{Var}[X_n].$$

Example 6.2 Problem

At a party of $n \geq 2$ people, each person throws a hat in a common box. The box is shaken and each person blindly draws a hat from the box without replacement. We say a match occurs if a person draws his own hat. What are the expected value and variance of V_n, the number of matches?

Example 6.2 Solution

Let X_i denote an indicator random variable such that

$$X_i = \begin{cases} 1 & \text{person } i \text{ draws his hat,} \\ 0 & \text{otherwise.} \end{cases}$$

The number of matches is $V_n = X_1 + \cdots + X_n$. Note that the X_i are generally not independent. For example, with $n = 2$ people, if the first person draws his own hat, then the second person must also draw her own hat. Note that the ith person is equally likely to draw any of the n hats, thus $P_{X_i}(1) = 1/n$ and $E[X_i] = P_{X_i}(1) = 1/n$. Since the expected value of the sum always equals the sum of the expected values,

$$E[V_n] = E[X_1] + \cdots + E[X_n] = n(1/n) = 1.$$

To find the variance of V_n, we will use Theorem 6.2. The variance of X_i is

$$\text{Var}[X_i] = E\left[X_i^2\right] - (E[X_i])^2 = \frac{1}{n} - \frac{1}{n^2}.$$

To find $\text{Cov}[X_i, X_j]$, we observe that

$$\text{Cov}\left[X_i, X_j\right] = E\left[X_i X_j\right] - E[X_i]\,E\left[X_j\right].$$

[Continued]

Example 6.2 Solution (continued)

Note that $X_iX_j = 1$ if and only if $X_i = 1$ and $X_j = 1$, and that $X_iX_j = 0$ otherwise. Thus

$$E\left[X_iX_j\right] = P_{X_i,X_j}(1,1) = P_{X_i|X_j}(1|1)\, P_{X_j}(1)\,.$$

Given $X_j = 1$, that is, the jth person drew his own hat, then $X_i = 1$ if and only if the ith person draws his own hat from the $n-1$ other hats. Hence $P_{X_i|X_j}(1|1) = 1/(n-1)$ and

$$E\left[X_iX_j\right] = \frac{1}{n(n-1)}, \qquad \mathrm{Cov}\left[X_i, X_j\right] = \frac{1}{n(n-1)} - \frac{1}{n^2}.$$

Finally, we can use Theorem 6.2 to calculate

$$\mathrm{Var}[V_n] = n\,\mathrm{Var}[X_i] + n(n-1)\,\mathrm{Cov}\left[X_i, X_j\right] = 1.$$

That is, both the expected value and variance of V_n are 1, no matter how large n is!

Example 6.3 Problem

Continuing Example 6.2, suppose each person immediately returns to the box the hat that he or she drew. What is the expected value and variance of V_n, the number of matches?

Example 6.3 Solution

In this case the indicator random variables X_i are iid because each person draws from the same bin containing all n hats. The number of matches $V_n = X_1 + \cdots + X_n$ is the sum of n iid random variables. As before, the expected value of V_n is

$$E[V_n] = nE[X_i] = 1.$$

In this case, the variance of V_n equals the sum of the variances,

$$\mathrm{Var}[V_n] = n\,\mathrm{Var}[X_i] = n\left(\frac{1}{n} - \frac{1}{n^2}\right) = 1 - \frac{1}{n}.$$

Transformations of Random Variable and Random Vector

1. Characteristic Functions

2. Moment Generating Functions

3. Probability Generating Functions

1. Characteristic Functions

The characteristic function of a random variable X is defined by

$$\Phi_X(\omega) \;=\; E\left\{ e^{j\omega X} \right\} \;=\; \int_{-\infty}^{\infty} f_X(x)\, e^{j\omega x}\, dx$$

$\Phi_X(\omega)$ may be viewed as the Fourier transform of the pdf $f_X(x)$ (with a reversal in the sign of the exponent).

In the case, using the reverse Fourier transform, we obtain

$$f_X(x) \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_X(\omega)\; e^{-j\omega x}\; d\omega.$$

Characteristic Function of the Gaussian R.V.

Let

$$f_X(x) \;=\; \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu_X)^2/2\sigma^2}.$$

By definition

$$\Phi_X(\omega) \;=\; \int_{-\infty}^{\infty} f_X(x)\, e^{j\omega x}\; dx$$

$$=\; \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\; \exp\left[\tfrac{-(x-\mu_X)^2}{2\sigma^2} + j\omega\, x\right] dx.$$

Let $x-\mu_X = u$, then

$$\Phi_X(\omega) \;=\; \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\; \exp\left[\tfrac{-u^2}{2\sigma^2} + j\omega\, u + j\omega\, m\right] du$$

$$=\; \frac{1}{\sqrt{2\pi}\sigma} e^{j\,\mu_X\,\omega} \int_{-\infty}^{\infty}\; \exp\left[\tfrac{-u^2}{2\sigma^2} + j\omega\, u\right] du$$

$$=\; \frac{1}{\sqrt{2\pi}\sigma} e^{j\,\mu_X\,\omega} \int_{-\infty}^{\infty}\; \exp\left[\tfrac{-(u-j\sigma^2\omega)^2}{2\sigma^2} - \tfrac{1}{2}\sigma^2\omega^2\right] du$$

$$=\; \exp\left(j\,\mu_X\;\omega - \tfrac{1}{2}\sigma^2\omega^2\right)\; \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{\frac{-(u-j\sigma^2\omega)^2}{2\sigma^2}}\, du.$$

Let $u - j\sigma^2\omega = y$, then

$$\Phi_X(\omega) = \exp\left(j\,\mu_X\,\omega - \tfrac{1}{2}\sigma^2\omega^2\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-y^2/2\sigma^2}\,dy$$

$$= \exp\left(j\,\mu_X\,\omega - \tfrac{1}{2}\sigma^2\omega^2\right).$$

Characteristic Function for Discrete Random Variable

For a discrete random variable x,

$$f_{\text{x}}(x) \;=\; \sum_{\text{i}} p_k \;\delta(x - x_k), \qquad p_k \;=\; \text{P}\;(\text{x} = x_k)$$

$\Phi_{\text{x}}(\omega)$ gives

$$\Phi_{\text{x}}(\omega) = \sum_k p_{\text{x}}(x_k) e^{j\omega x_k}.$$

In particular, if x_k are integer-valued, the characteristic function is then

$$\Phi_{\text{x}}(\omega) = \sum_{k=-\infty}^{\infty} p_{\text{x}}(k) e^{j\omega k}. \qquad (*)$$

Eq. (*) is the Fourier transform of the sequence $p_{\text{x}}(k)$. Note that Eq. (*) is a periodic function of ω with period 2π since $e^{j(\omega+2\pi)k} = e^{j\omega k} e^{j2\pi k} = e^{j\omega k}$.

Thus, the following inversion formula allows us to recover the probabilities $p_{\text{x}}(k)$ from $\Phi_{\text{x}}(\omega)$:

$$p_{\text{x}}(k) = \frac{1}{2\pi} \int_0^{2\pi} \Phi_{\text{x}}(\omega) e^{-j\omega k} d\omega, \quad k = 0, \pm 1, \pm 2, \cdots \qquad (**)$$

A comparison of Eqs. (*) and (**) shows that the $p_{\text{x}}(k)$ are the coefficients of the Fourier series of the periodic function $\Phi_{\text{x}}(\omega)$.

Properties of the Characteristic Function

1st Property:

$$
\begin{aligned}
\left|\Phi_X(\omega)\right| &= \left| E\left\{ e^{j\,\omega\,x} \right\} \right| \\
&\le E\left\{ \left| e^{j\,\omega\,x} \right| \right\} \\
&= E\{\,1\,\} \\
&= 1
\end{aligned}
$$

$$
\left|\Phi_X(\omega)\right| \le \Phi_X(0) = 1.
$$

Properties of the Characteristic Function (continued)

2nd property: moment generating property of $\Phi_X(\omega)$

$$\frac{d\,\Phi_X(\omega)}{d\omega} = j\int_{-\infty}^{\infty} x\, f_X(x)\, e^{j\omega x}\, dx$$

At $\omega = 0$,

$$\frac{d\,\Phi_X(0)}{d\omega} = j\int_{-\infty}^{\infty} x\, f_X(x)\; dx = j\,E\{X\}$$

$$E\{X\} = -j\,\frac{d\,\Phi_X(0)}{d\omega}$$

In genereal,

$$E\{X^n\} = m_n = (-j)^n \left.\frac{d^n\,\Phi_X(0)}{d\omega^n}\right|_{\omega=0}.$$

If $\Phi_X(\omega)$ is analytic around $\omega = 0$,

$$\Phi_X(\omega) = \sum_{k=0}^{\infty} \Phi_X^{(k)}(0)\,\frac{\omega^k}{k\,!}$$

$$= \sum_{k=0}^{\infty} E\{X^k\}\,\frac{(j\omega)^k}{k\,!}.$$

$$\Phi_X(\omega) = E[e^{j\omega X}]$$

$$= E\left[e^{j0X} + \frac{de^{j0X}}{d\omega}\omega + \frac{1}{2!}\frac{d^2e^{j0X}}{d\omega^2}\omega^2 + \cdots\right]$$

$$= E[e^{j0X}] + E\left[\frac{de^{j0X}}{d\omega}\omega\right] + E\left[\frac{1}{2!}\frac{d^2e^{j0X}}{d\omega^2}\omega^2\right] + \cdots$$

Moment Theorem

If the characteristic function $\Phi_X(\omega)$ of a given random variable x has a Taylor series expansion which is valid in some interval in ω which contains the origin, then that characteristic function (and hence the corresponding probability density of probability distribution) is uniquely determined by the moments of the given random variable.

Section 6.3

Moment Generating Functions

Definition 6.1 *(MGF)*

For a random variable X, the moment generating function (MGF) *of X is*

$$\phi_X(s) = E\left[e^{sX}\right].$$

Table 6.1

See the text.

Example (Moment Generating Function of the Unitary Gaussian)

Suppose $X = N(0,1)$, then

$$M_X(\nu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\nu x} e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\nu x - x^2/2} dx$$

(From an integral table, we see that $\int_{-\infty}^{\infty} e^{-a^2x^2+bx} dx = \frac{\sqrt{\pi}}{a} e^{b^2/(4a^2)}$.

Apply $a = \frac{1}{\sqrt{2}}, b = \nu$)

$$= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi}\ e^{\nu^2/2} = e^{\nu^2/2}.$$

Theorem 6.6

A random variable X with MGF $\phi_X(s)$ has nth moment

$$E\left[X^n\right] = \left.\frac{d^n \phi_X(s)}{ds^n}\right|_{s=0}.$$

Proof: Theorem 6.6

The first derivative of $\phi_X(s)$ is

$$\frac{d\phi_X(s)}{ds} = \frac{d}{ds}\left(\int_{-\infty}^{\infty} e^{sx} f_X(x)\, dx\right) = \int_{-\infty}^{\infty} x e^{sx} f_X(x)\, dx.$$

Evaluating this derivative at $s = 0$ proves the theorem for $n = 1$.

$$\left.\frac{d\phi_X(s)}{ds}\right|_{s=0} = \int_{-\infty}^{\infty} x f_X(x)\, dx = E[X].$$

Similarly, the nth derivative of $\phi_X(s)$ is

$$\frac{d^n\phi_X(s)}{ds^n} = \int_{-\infty}^{\infty} x^n e^{sx} f_X(x)\, dx.$$

The integral evaluated at $s = 0$ is the formula in the theorem statement.

Example 6.5 Problem

exponential r.v.

X is an exponential r.v. with MGF $\phi_X(s) = \lambda/(\lambda - s)$. What are the first and second moments of X? Write a general expression for the nth moment.

Example 6.5 Solution

The first moment is the expected value:

$$E[X] = \left.\frac{d\phi_X(s)}{ds}\right|_{s=0} = \left.\frac{\lambda}{(\lambda - s)^2}\right|_{s=0} = \frac{1}{\lambda}.$$

The second moment of X is the mean square value:

$$E\left[X^2\right] = \left.\frac{d^2\phi_X(s)}{ds^2}\right|_{s=0} = \left.\frac{2\lambda}{(\lambda - s)^3}\right|_{s=0} = \frac{2}{\lambda^2}.$$

Proceeding in this way, it should become apparent that the nth moment of X is

$$E[X^n] = \left.\frac{d^n\phi_X(s)}{ds^n}\right|_{s=0} = \left.\frac{n!\lambda}{(\lambda - s)^{n+1}}\right|_{s=0} = \frac{n!}{\lambda^n}.$$

Theorem 6.7

The MGF of $Y = aX + b$ is $\phi_Y(s) = e^{sb}\phi_X(as)$.

Proof: Theorem 6.7

From the definition of the MGF,

$$\phi_Y(s) = E\left[e^{s(aX+b)}\right] = e^{sb}E\left[e^{(as)X}\right] = e^{sb}\phi_X(as).$$

3. Probability Generating Function for Discrete R.V.

The probability generating function $G_N(z)$ of a nonnegative integer-valued random variable N is defined by

$$G_N(z) = E\left[z^N\right] = \sum_{k=0}^{\infty} p_N(k)\, z^k$$

Using a derivation similar to that used in the moment property, we can show that the PMF of N is given by

$$p_N(k) = \frac{1}{k\,!} \frac{d^k}{dz^k} G_N(z)\bigg|_{z=1} .$$

This is why $G_N(z)$ is called the probability generating function.

By taking the first two derivatives of $G_N(z)$ and evaluating the result at z = 1, it is possible to find the first two moments of N:

$$\frac{d}{dz}G_N(z)\bigg|_{z=1} = \sum_{\mathrm{k}=0}^{\infty} p_N(k)\ k\ z^{k-1}\bigg|_{z=1} = \sum_{\mathrm{k}=0}^{\infty} k\ p_N(k) = E[N]$$

and

$$\frac{d^2}{dz^2}G_N(z)\bigg|_{z=1} = \sum_{\mathrm{k}=0}^{\infty} p_N(k)\ k(k-1)\ z^{k-2}\bigg|_{z=1}$$

$$= \sum_{\mathrm{k}=0}^{\infty} k(k-1)\ p_N(k)$$

$$= E[N(N-1)] = E[N^2] - E[N]$$

$$VAR\ [N] = E[N^2] - \{\ E[N]\ \}^2$$

$$= G_N''(1) + G_N'(1) - [\ G_N'(1)\ ]^2$$

Example (PGF for Poisson R.V.)

$$G_N(z) = \sum_{k=0}^{\infty} \frac{\alpha^k}{k\,!} e^{-\alpha} z^k$$

$$= e^{-\alpha} \sum_{k=0}^{\infty} \frac{(\alpha z)^k}{k\,!}$$

$$= e^{-\alpha} e^{\alpha z} = e^{\alpha (z-1)}$$

$$G'_N(z) = \alpha e^{\alpha (z-1)}$$

$$G''_N(z) = \alpha^2 e^{\alpha (z-1)}$$

$$E[N] = G'_N(1) = \alpha$$

$$VAR[N] = \alpha^2 + \alpha - \alpha^2 = \alpha$$

Section 6.4

MGF of the Sum of Independent Random Variables

Theorem 6.8

For a set of independent random variables $X_1, \ldots, X_n$, the moment generating function of $W = X_1 + \cdots + X_n$ is

$$\phi_W(s) = \phi_{X_1}(s)\phi_{X_2}(s) \cdots \phi_{X_n}(s).$$

When $X_1, \ldots, X_n$ are iid, each with MGF $\phi_{X_i}(s) = \phi_X(s)$,

$$\phi_W(s) = [\phi_X(s)]^n .$$

Similarly, the following hold:

$$\Phi_W(\omega) = E\{e^{j\omega(X_1+X_2+\cdots+X_n)}\} = E\{e^{j\omega X_1}\}\cdots E\{e^{j\omega X_n}\} = [\Phi_X(\omega)]^n$$

$$G_W(z) = E\{z^W\} = E\{z^{X_1+X_2+\cdots+X_n}\} = E\{z^{X_1}\}\cdots E\{z^{X_n}\} = G_{X_1}(z)\cdots G_{X_n}(z) = [G_X(z)]^n.$$

Proof: Theorem 6.8

From the definition of the MGF,

$$\phi_W(s) = E\left[e^{s(X_1+\cdots+X_n)}\right] = E\left[e^{sX_1}e^{sX_2}\cdots e^{sX_n}\right].$$

Here, we have the expected value of a product of functions of independent random variables. Theorem 5.9 states that this expected value is the product of the individual expected values:

$$E\left[g_1(X_1)g_2(X_2)\cdots g_n(X_n)\right] = E\left[g_1(X_1)\right]E\left[g_2(X_2)\right]\cdots E\left[g_n(X_n)\right].$$

By Equation (6.38) with $g_i(X_i) = e^{sX_i}$, the expected value of the product is

$$\phi_W(s) = E\left[e^{sX_1}\right]E\left[e^{sX_2}\right]\cdots E\left[e^{sX_n}\right] = \phi_{X_1}(s)\phi_{X_2}(s)\cdots\phi_{X_n}(s).$$

When $X_1, \ldots, X_n$ are iid, $\phi_{X_i}(s) = \phi_X(s)$ and thus $\phi_W(s) = (\phi_W(s))^n$.

Theorem 6.9

If $K_1, \ldots, K_n$ are independent Poisson random variables, $W = K_1 + \cdots + K_n$ is a Poisson random variable.

Proof: Theorem 6.9

We adopt the notation $E[K_i] = \alpha_i$ and note in Table 6.1 that K_i has MGF $\phi_{K_i}(s) = e^{\alpha_i(e^s-1)}$. By Theorem 6.8,

$$\phi_W(s) = e^{\alpha_1(e^s-1)} e^{\alpha_2(e^s-1)} \cdots e^{\alpha_n(e^s-1)} = e^{(\alpha_1+\cdots+\alpha_n)(e^s-1)} = e^{(\alpha_T)(e^s-1)}$$

where $\alpha_T = \alpha_1 + \cdots + \alpha_n$. Examining Table 6.1, we observe that $\phi_W(s)$ is the moment generating function of the Poisson (α_T) random variable. Therefore,

$$P_W(w) = \begin{cases} \alpha_T^w e^{-\alpha}/w! & w = 0, 1, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 6.10

The sum of n independent Gaussian random variables $W = X_1 + \cdots + X_n$ is a Gaussian random variable.

Proof: Theorem 6.10

For convenience, let $\mu_i = E[X_i]$ and $\sigma_i^2 = \text{Var}[X_i]$. Since the X_i are independent, we know that

$$\begin{aligned}\phi_W(s) &= \phi_{X_1}(s)\phi_{X_2}(s)\cdots\phi_{X_n}(s) \\ &= e^{s\mu_1+\sigma_1^2 s^2/2} e^{s\mu_2+\sigma_2^2 s^2/2} \cdots e^{s\mu_n+\sigma_n^2 s^2/2} \\ &= e^{s(\mu_1+\cdots+\mu_n)+(\sigma_1^2+\cdots+\sigma_n^2)s^2/2}.\end{aligned}$$

From Equation (6.51), we observe that $\phi_W(s)$ is the moment generating function of a Gaussian random variable with expected value $\mu_1 + \cdots + \mu_n$ and variance $\sigma_1^2 + \cdots + \sigma_n^2$.

Theorem 6.11

If $X_1, \ldots, X_n$ are iid exponential (λ) random variables, then $W = X_1 + \cdots + X_n$ has the Erlang PDF

$$f_W(w) = \begin{cases} \frac{\lambda^n w^{n-1} e^{-\lambda w}}{(n-1)!} & w \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof: Theorem 6.11

In Table 6.1 we observe that each X_i has MGF $\phi_X(s) = \lambda/(\lambda - s)$. By Theorem 6.8, W has MGF

$$\phi_W(s) = \left(\frac{\lambda}{\lambda - s}\right)^n .$$

Returning to Table 6.1, we see that W has the MGF of an Erlang (n, λ) random variable.

Quiz 6.4(A)

Let $K_1, K_2, \ldots, K_m$ be iid discrete uniform random variables with PMF

$$P_K(k) = \begin{cases} 1/n & k = 1, 2, \ldots, n, \\ 0 & \text{otherwise.} \end{cases}$$

Find the MGF of $J = K_1 + \cdots + K_m$.

Quiz 6.4(A) Solution

Each K_i has MGF

$$\phi_K(s) = E\left[e^{sK_i}\right] = \frac{e^s + e^{2s} + \cdots + e^{ns}}{n} = \frac{e^s(1-e^{ns})}{n(1-e^s)}$$

Since the sequence of K_i is independent, Theorem 6.8 says the MGF of J is

$$\phi_J(s) = (\phi_K(s))^m = \frac{e^{ms}(1-e^{ns})^m}{n^m(1-e^s)^m}$$

Section 6.5

Random Sums of Independent Random Variables

Theorem 6.12

Let $\{X_1, X_2, \ldots\}$ be a collection of iid random variables, each with MGF $\phi_X(s)$, and let N be a nonnegative integer-valued random variable that is independent of $\{X_1, X_2, \ldots\}$. The random sum $R = X_1 + \cdots + X_N$ has moment generating function

$$\phi_R(s) = \phi_N(\ln \phi_X(s)).$$

Proof: Theorem 6.12

To find $\phi_R(s) = E[e^{sR}]$, we first find the conditional expected value $E[e^{sR}|N = n]$. Because this expected value is a function of n, it is a random variable. Theorem 4.26 states that $\phi_R(s)$ is the expected value, with respect to N, of $E[e^{sR}|N = n]$:

$$\phi_R(s) = \sum_{n=0}^{\infty} E\left[e^{sR}|N = n\right] P_N(n) = \sum_{n=0}^{\infty} E\left[e^{s(X_1+\cdots+X_N)}|N = n\right] P_N(n).$$

Because the X_i are independent of N,

$$E\left[e^{s(X_1+\cdots+X_N)}|N = n\right] = E\left[e^{s(X_1+\cdots+X_n)}\right] = E\left[e^{sW}\right] = \phi_W(s).$$

In Equation (6.58), $W = X_1 + \cdots + X_n$. From Theorem 6.8, we know that $\phi_W(s) = [\phi_X(s)]^n$, implying

$$\phi_R(s) = \sum_{n=0}^{\infty} [\phi_X(s)]^n P_N(n).$$

We observe that we can write $[\phi_X(s)]^n = [e^{\ln \phi_X(s)}]^n = e^{[\ln \phi_X(s)]n}$. This implies

$$\phi_R(s) = \sum_{n=0}^{\infty} e^{[\ln \phi_X(s)]n} P_N(n).$$

Discrete MGF

Recognizing that this sum has the same form as the sum in Equation (6.27), we infer that the sum is $\phi_N(s)$ evaluated at $s = \ln \phi_X(s)$. Therefore, $\phi_R(s) = \phi_N(\ln \phi_X(s))$.

An Alternative Using the Characteristic Function and the Probability Generating Function

Find the characteristic function of S_N defined by

$$S_N = \sum_{k=1}^{N} X_k$$

where

N, X_k's = random variables (N and X_k are independent.)

X_k's = iid r.v.'s.

(Solution)

$$E\{e^{j\omega S_N} \mid N = n\} = E\{e^{j\omega(X_1+\cdots+X_n)}\} = [\Phi_X(\omega)]^n$$

or

$$E\{e^{j\omega S_N} \mid N\} = [\Phi_X(\omega)]^N.$$

The characteristic function of S_N is given by

$$\Phi_{S_N}(\omega) = E\{E\{e^{j\omega S_N} \mid N\}\} = E\{[\Phi_X(\omega)]^N\} = E\{Z^N\}|_{Z=\Phi_X(\omega)} = G_N(\Phi_X(\omega)).$$

Example (An Alternative to Quiz 6.5)

N = the number of jobs submitted to a computer in an hour;

a geometic random variable with parameter p.

X = the job execution times; independent exponentially distributed random variables with mean $1/\alpha$.

Find the PDF for the sum of the execution times of the jobs submitted in an hour $R = X_1 + \cdots + X_N$.

(Solution 1 Using MGF)

$$\phi_X(s) = \frac{\lambda}{\lambda - s}, \quad \phi_N(s) = \frac{pe^s}{1-(1-p)e^s}.$$

From Theorem 6.12, R has MGF

$$\phi_R(s) = \phi_N\left(\ln \phi_X(s)\right) = \frac{p\phi_X(s)}{1-(1-p)\phi_X(s)} = \frac{p\lambda}{p\lambda - s}.$$

The corresponding PDF is

$$f_R(r) = \begin{cases} (p\lambda)e^{-p\lambda r} & r \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

(Solution 2 Using CF and PGF)

$$\Phi_X(\omega) = \frac{\lambda}{\lambda - j\omega}, \quad G_N(z) = \frac{pz}{1-(1-p)z}.$$

From Theorem 6.12, R has MGF

$$\Phi_R(s) = G_N\left(\Phi_X(\omega)\right) = \frac{p\Phi_X(\omega)}{1-(1-p)\Phi_X(\omega)} = \frac{p\lambda}{p\lambda - j\omega}.$$

The corresponding PDF is

$$f_R(r) = \begin{cases} (p\lambda)e^{-p\lambda r} & r \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Exercise: Repeat the problem for $R = X_0 + X_1 + \cdots + X_N$.

Example 6.9 Problem

The number of pages N in a fax transmission has a geometric PMF with expected value $1/q = 4$. The number of bits K in a fax page also has a geometric distribution with expected value $1/p = 10^5$ bits, independent of the number of bits in any other page and independent of the number of pages. Find the MGF and the PMF of B, the total number of bits in a fax transmission.

Example 6.9 Solution

When the ith page has K_i bits, the total number of bits is the random sum $B = K_1 + \cdots + K_N$. Thus $\phi_B(s) = \phi_N(\ln \phi_K(s))$. From Table 6.1,

$$\phi_N(s) = \frac{qe^s}{1-(1-q)e^s}, \qquad \phi_K(s) = \frac{pe^s}{1-(1-p)e^s}.$$

To calculate $\phi_B(s)$, we substitute $\ln \phi_K(s)$ for every occurrence of s in $\phi_N(s)$. Equivalently, we can substitute $\phi_K(s)$ for every occurrence of e^s in $\phi_N(s)$. This substitution yields

$$\phi_B(s) = \frac{q\left(\frac{pe^s}{1-(1-p)e^s}\right)}{1-(1-q)\left(\frac{pe^s}{1-(1-p)e^s}\right)} = \frac{pqe^s}{1-(1-pq)e^s}.$$

By comparing $\phi_K(s)$ and $\phi_B(s)$, we see that B has the MGF of a geometric $(pq = 2.5 \times 10^{-5})$ random variable with expected value $1/(pq) = 400{,}000$ bits. Therefore, B has the geometric PMF

$$P_B(b) = \begin{cases} pq(1-pq)^{b-1} & b = 1, 2, \ldots, \\ 0 & \text{otherwise,} \end{cases}$$

Theorem 6.13

random sum

For the ^ of iid random variables $R = X_1 + \cdots + X_N$,

$$E[R] = E[N]E[X], \qquad \text{Var}[R] = E[N]\text{Var}[X] + \text{Var}[N](E[X])^2.$$

Proof: Theorem 6.13

By the chain rule for derivatives,

$$\phi_R'(s) = \phi_N'(\ln \phi_X(s)) \frac{\phi_X'(s)}{\phi_X(s)}.$$

Since $\phi_X(0) = 1$, $\phi_N'(0) = E[N]$, and $\phi_X'(0) = E[X]$, evaluating the equation at $s = 0$ yields

$$E[R] = \phi_R'(0) = \phi_N'(0) \frac{\phi_X'(0)}{\phi_X(0)} = E[N]\, E[X].$$

For the second derivative of $\phi_X(s)$, we have

$$\phi_R''(s) = \phi_N''(\ln \phi_X(s)) \left(\frac{\phi_X'(s)}{\phi_X(s)} \right)^2 + \phi_N'(\ln \phi_X(s)) \frac{\phi_X(s)\phi_X''(s) - \left[\phi_X'(s)\right]^2}{[\phi_X(s)]^2}.$$

The value of this derivative at $s = 0$ is

$$E\left[R^2\right] = E\left[N^2\right] \mu_X^2 + E[N] \left(E\left[X^2\right] - \mu_X^2 \right).$$

Subtracting $(E[R])^2 = (\mu_N \mu_X)^2$ from both sides of this equation completes the proof.

Example 6.10 Problem

Let $X_1, X_2 \ldots$ be a sequence of independent Gaussian (100,10) random variables. If K is a Poisson (1) random variable independent of $X_1, X_2 \ldots$, find the expected value and variance of $R = X_1 + \cdots + X_K$.

Example 6.10 Solution

The PDF and MGF of R are complicated. However, Theorem 6.13 simplifies the calculation of the expected value and the variance. From Appendix A, we observe that a Poisson (1) random variable also has variance 1. Thus

$$E[R] = E[X]\,E[K] = 100,$$

and

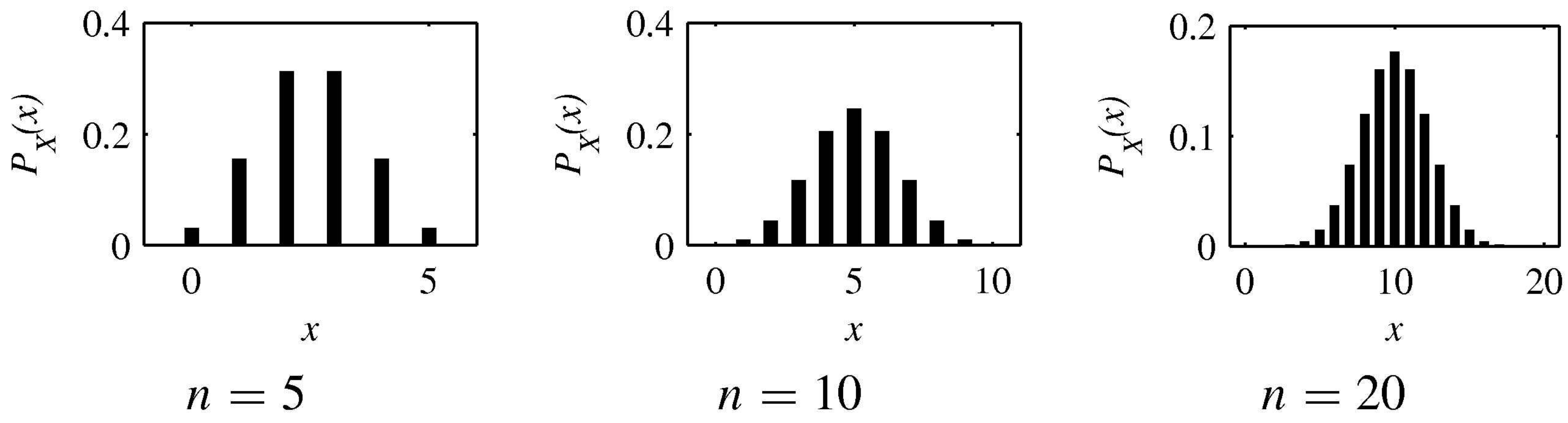
$$\text{Var}[R] = E[K]\,\text{Var}[X] + \text{Var}[K]\,(E[X])^2 = 100 + (100)^2 = 10,100.$$

We see that most of the variance is contributed by the randomness in K. This is true because K is very likely to take on the values 0 and 1, and those two choices dramatically affect the sum.

Section 6.6

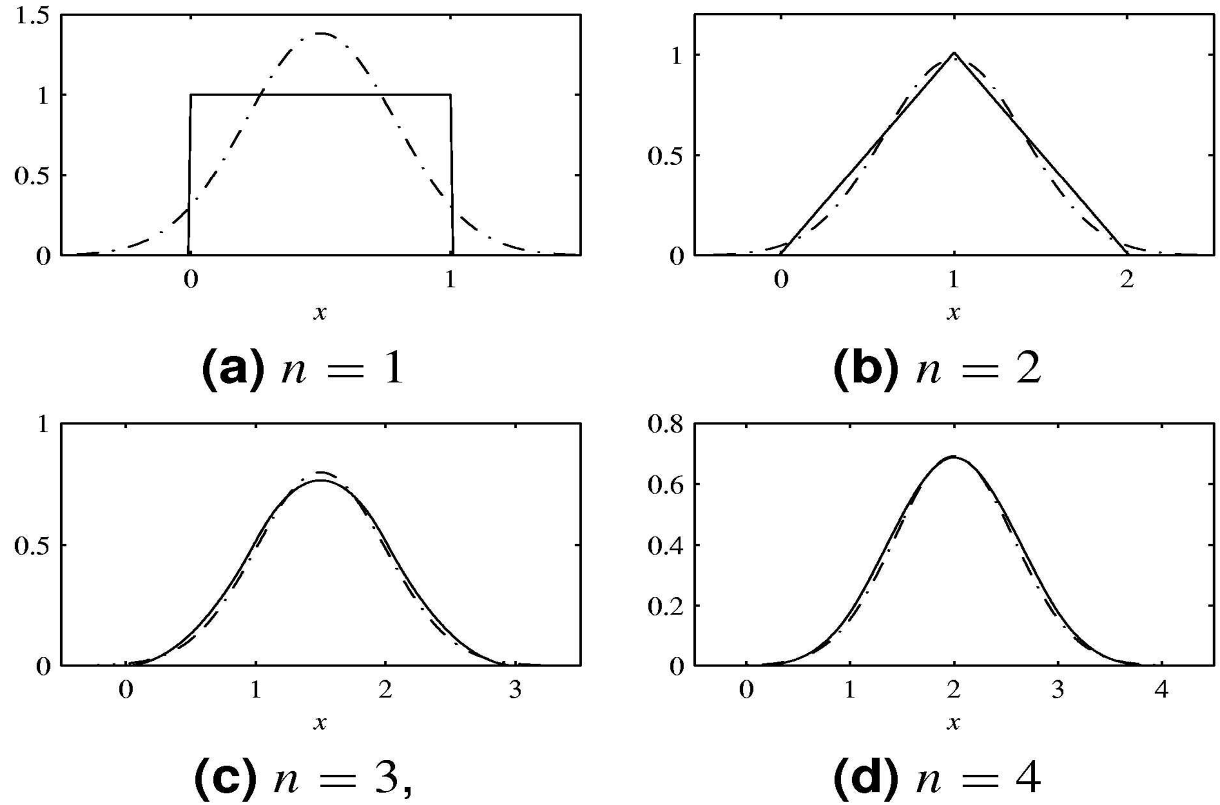
Central Limit Theorem

Figure 6.1



The PMF of the X, the number of heads in n coin flips for $n = 5, 10, 20$. As n increases, the PMF more closely resembles a bell-shaped curve.

Figure 6.2



The PDF of W_n, the sum of n uniform $(0, 1)$ random variables, and the corresponding central limit theorem approximation for *n*=1,2,3,4. The solid line denotes the PDF $f_{W_n}(w)$ while the dotted line denotes the Gaussian approximation.

Central Limit Theorem

Let's revisit the moment generating function which was defined

$$\phi_x(t) = E\left\{e^{tx}\right\} = \int_{-\infty}^{\infty} f_x(x)e^{tx}dx$$

Suppose x = $N(0,1)$, then

$$\phi_x(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx}e^{-x^2/2}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx-x^2/2}dx$$

(From an integral table, we see $\int_{-\infty}^{\infty} e^{-a^2x^2+bx}dx = \frac{\sqrt{\pi}}{a}e^{b^2/(4a^2)}$. Let $a = \frac{1}{\sqrt{2}}, b = t$.)

$$= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi}\ e^{t^2/2} = e^{\frac{t^2}{2}}. \qquad (*)$$

Theorem

Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed (iid) random variables each having mean η and variances σ^2.

Then the distribution of

$$\frac{X_1 + X_2 + \cdots + X_n - n\eta}{\sigma\sqrt{n}}$$

tends to the standard normal as $n \to \infty$.

That is, $P\left\{\frac{X_1 + X_2 + \cdots + X_n - n\eta}{\sigma\sqrt{n}} \leq a\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-x^2/2}dx$ as $n \to \infty$.

Proof

Let $\eta = 0,\ \sigma^2 = 1$ for convenience. The moment generating function of $X_i \big/ \sqrt{n}$ is given by

$$\phi_{X_i/\sqrt{n}}(t) = E\left\{ \exp\left[\frac{t\,X_i}{\sqrt{n}}\right]\right\} = \phi_{\mathrm{X}}\left(\frac{t}{\sqrt{n}}\right).$$

The moment generating function of $\sum_{i=1}^{n} X_i \big/ \sqrt{n}$ is given by

$$\phi_{\sum_{i=1}^{n} X_i/\sqrt{n}}(t) = \left[\phi_{\mathrm{X}}\left(\frac{t}{\sqrt{n}}\right)\right]^n .$$

The Taylor series of $\phi_X(t)$ around t=0 is given by

$$\begin{aligned}
\phi_X(t) &= E\left\{ e^{tX} \right\} = \phi_X(0) + \phi_X{}'(0)\,t + \phi_X{}''(0)\,t^2 \big/ 2 + O(t^2) \\
&= 1 + t\,E\{X\} + \frac{t^2 E\{X^2\}}{2} + O(t^2) \\
&= 1 + \frac{t^2}{2} + O(t^2). \qquad \left(E\{X\} = 0, E\{X^2\} = 1 \right)
\end{aligned}$$

Applied was the moment generating property:

$$\phi_X{}'(0) = {d\,E\{e^{tX}\}}\Big/{dt}\Big|_{t=0} = E\{\,X\,e^{tX}\,\}\Big|_{t=0} = E\{X\}$$

$$\phi_X{}''(0) = E\left\{X^2\right\}.$$

Therefore,

$$\phi_X\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{t^2}{2n} + O(t^2)$$

$$\phi_{\sum_{i=1}^{n} X_i/\sqrt{n}}(t) = \left[\,1 + \frac{t^2}{2n} + O(t^2)\,\right]^n$$

Referring to Eq. (*), we need to show

$$\left[\,1 + \frac{t^2}{2n} + O(t^2)\,\right]^n \to e^{t^2/2} \;\; as \quad n \to \infty.$$

According to L'Hospital rule

$$\log\left(1+x\right)=x+O(x^2)\quad,\quad x<<1.$$

(Note that $\lim_{x\to 0}\left[\frac{\log\left(1+x\right)-x}{x}\right]=\lim_{x\to 0}\left[\frac{\frac{1}{1+x}-1}{1}\right]=0.$)

Now,

$$\log\left[1+\frac{t^2}{2n}+O\left(\frac{t^2}{n}\right)\right]^n=n\log\left[1+\frac{t^2}{2n}+O\left(\frac{t^2}{n}\right)\right]$$

$$=n\left[\frac{t^2}{2n}+O\left(\frac{t^2}{n}\right)+O\left(\frac{t^2}{2n}+O\left(\frac{t^2}{n}\right)^2\right)\right]=\frac{t^2}{2}+n\,O\left(\frac{t^2}{n}\right)+n\,O\left(\frac{t^2}{2n}\right).$$

Applying that

$$n\,O\left(\frac{t^2}{n}\right)=t^2\,\frac{O\left(t^2/n\right)}{t^2/n}\rightarrow 0 \qquad as \quad n\rightarrow\infty$$

$$n\,O\left(\frac{t^2}{2n}\right)=t^2\,\frac{O\left(t^2/2n\right)}{t^2/2n}\rightarrow 0 \qquad as \quad n\rightarrow\infty$$

we have

$$\log\left[\;1+\frac{t^2}{2n}+O\left(\frac{t^2}{n}\right)\;\right]^n\rightarrow\frac{t^2}{2} \qquad as \quad n\rightarrow\infty.$$

That is,

$$\left[\;1+\frac{t^2}{2n}+O\left(\frac{t^2}{n}\right)\;\right]^n\rightarrow e^{t^2/2} \quad as \quad n\rightarrow\infty.$$

For arbitrary η and σ^2, define the standardized r.v.

$$X_i^* = (\,X_i-\eta\,)\,/\sigma.$$

Then, we can obtain the same results since $E\{\,X_i^*\,\}=0$ and $\mathrm{var}\{\,X_i^*\,\}=1$.

Section 6.7

Applications of the Central Limit Theorem

Example 6.13 Problem

A compact disc (CD) contains digitized samples of an acoustic waveform. In a CD player with a "one bit digital to analog converter," each digital sample is represented to an accuracy of ± 0.5 mV. The CD player "oversamples" the waveform by making eight independent measurements corresponding to each sample. The CD player obtains a waveform sample by calculating the average (sample mean) of the eight measurements. What is the probability that the error in the waveform sample is greater than 0.1 mV?

Example 6.13 Solution

The measurements $X_1, X_2, \ldots, X_8$ all have a uniform distribution between $v - 0.5$ mV and $v + 0.5$ mV, where v mV is the exact value of the waveform sample. The compact disk player produces the output $U = W_8/8$, where

$$W_8 = \sum_{i=1}^{8} X_i.$$

To find $P[|U - v| > 0.1]$ exactly, we would have to find an exact probability model for W_8, either by computing an eightfold convolution of the uniform PDF of X_i or by using the moment generating function. Either way, the process is extremely complex. Alternatively, we can use the central limit theorem to model W_8 as a Gaussian random variable with $E[W_8] = 8\mu_X = 8v$ mV and variance $\text{Var}[W_8] = 8\,\text{Var}[X] = 8/12$. Therefore, U is approximately Gaussian with $E[U] = E[W_8]/8 = v$ and variance $\text{Var}[W_8]/64 = 1/96$. Finally, the error, $U - v$ in the output waveform sample is approximately Gaussian with expected value 0 and variance $1/96$. It follows that

$$P[|U - v| > 0.1] = 2\left[1 - \Phi\left(0.1/\sqrt{1/96}\right)\right] = 0.3272.$$

For a Gaussian X,

$$P[a < X \leq b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Example 6.14 Problem

A modem transmits one million bits. Each bit is 0 or 1 independently with equal probability. Estimate the probability of at least 502,000 ones.

Example 6.14 Solution

Let X_i be the value of bit i (either 0 or 1). The number of ones in one million bits is $W = \sum_{i=1}^{10^6} X_i$. Because X_i is a Bernoulli (0.5) random variable, $E[X_i] = 0.5$ and $\text{Var}[X_i] = 0.25$ for all i. Note that $E[W] = 10^6 E[X_i] = 500{,}000$ and $\text{Var}[W] = 10^6 \,\text{Var}[X_i] = 250{,}000$. Therefore, $\sigma_W = 500$. By the central limit theorem approximation,

$$
\begin{aligned}
P[W \geq 502{,}000] &= 1 - P[W \leq 502{,}000] \\
&\approx 1 - \Phi\left(\frac{502{,}000 - 500{,}000}{500}\right) = 1 - \Phi(4).
\end{aligned}
$$

Using Table 3.1, we observe that $1 - \Phi(4) = Q(4) = 3.17 \times 10^{-5}$.

Definition 6.3 De Moivre–Laplace Formula

For a binomial (n, p) random variable K,

$$P[k_1 \le K \le k_2] \approx \Phi\left(\frac{k_2 + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - 0.5 - np}{\sqrt{np(1-p)}}\right).$$

Assumed are:

(1) n is large,

(2) $npq >> 1$, i.e., $p \approx q$,

(3) $|K - np|$ is the order of $\sqrt{npq}$.

Example 6.16 Problem

Let K be a binomial $(n = 20, p = 0.4)$ random variable. What is $P[K = 8]$?

Example 6.16 Solution

Since $E[K] = np = 8$ and $\mathrm{Var}[K] = np(1-p) = 4.8$, the central limit theorem approximation to K is a Gaussian random variable X with $E[X] = 8$ and $\mathrm{Var}[X] = 4.8$. Because X is a continuous random variable, $P[X = 8] = 0$, a useless approximation to $P[K = 8]$. On the other hand, the De Moivre–Laplace formula produces

$$
\begin{aligned}
P[8 \leq K \leq 8] &\approx P[7.5 \leq X \leq 8.5] \\
&= \Phi\left(\frac{0.5}{\sqrt{4.8}}\right) - \Phi\left(\frac{-0.5}{\sqrt{4.8}}\right) = 0.1803.
\end{aligned}
$$

The exact value is $\binom{20}{8}(0.4)^8(1-0.4)^{12} = 0.1797$.

Quiz 6.7

Telephone calls can be classified as voice (V) if someone is speaking or data (D) if there is a modem or fax transmission. Based on a lot of observations taken by the telephone company, we have the following probability model: $P[V] = 3/4$, $P[D] = 1/4$. Data calls and voice calls occur independently of one another. The random variable K_n is the number of voice calls in a collection of n phone calls.

(1) What is $E[K_{48}]$, the expected number of voice calls in a set of 48 calls?

(2) What is $\sigma_{K_{48}}$, the standard deviation of the number of voice calls in a set of 48 calls?

(3) Use the central limit theorem to estimate $P[30 \leq K_{48} \leq 42]$, the probability of between 30 and 42 voice calls in a set of 48 calls.

(4) Use the De Moivre–Laplace formula to estimate $P[30 \leq K_{48} \leq 42]$.

Quiz 6.7 Solution

Random variable K_n has a binomial distribution for n trials and success probability $P[V] = 3/4$.

(1) The expected number of voice calls out of 48 calls is $E[K_{48}] = 48P[V] = 36$.

(2) The variance of K_{48} is

$$\text{Var}[K_{48}] = 48P[V](1 - P[V]) = 48(3/4)(1/4) = 9$$

Thus K_{48} has standard deviation $\sigma_{K_{48}} = 3$.

(3) Using the ordinary central limit theorem and Table 3.1 yields

$$P[30 \leq K_{48} \leq 42] \approx \Phi\left(\frac{42-36}{3}\right) - \Phi\left(\frac{30-36}{3}\right) = \Phi(2) - \Phi(-2)$$

Recalling that $\Phi(-x) = 1 - \Phi(x)$, we have

$$P[30 \leq K_{48} \leq 42] \approx 2\Phi(2) - 1 = 0.9545$$

(4) Since K_{48} is a discrete random variable, we can use the De Moivre-Laplace approximation to estimate

$$P[30 \leq K_{48} \leq 42] \approx \Phi\left(\frac{42+0.5-36}{3}\right) - \Phi\left(\frac{30-0.5-36}{3}\right)$$
$$= 2\Phi(2.16666) - 1 = 0.9687$$

Section 6.8

The Chernoff Bound

Theorem 6.15 Chernoff Bound

For an arbitrary random variable X and a constant c,

$$P[X \geq c] \leq \min_{s \geq 0} e^{-sc} \phi_X(s).$$

Proof: Theorem 6.15

In terms of the unit step function, $u(x)$, we observe that

$$P[X \geq c] = \int_c^{\infty} f_X(x)\, dx = \int_{-\infty}^{\infty} u(x-c) f_X(x)\, dx.$$

For all $s \geq 0$, $u(x-c) \leq e^{s(x-c)}$. This implies

$$P[X \geq c] \leq \int_{-\infty}^{\infty} e^{s(x-c)} f_X(x)\, dx = e^{-sc} \int_{-\infty}^{\infty} e^{sx} f_X(x)\, dx = e^{-sc} \phi_X(s).$$

This inequality is true for any $s \geq 0$. Hence the upper bound must hold when we choose s to minimize $e^{-sc}\phi_X(s)$.

Example 6.18 Problem

If the height X, measured in feet, of a randomly chosen adult is a Gaussian (5.5, 1) random variable, use the Chernoff bound to find an upper bound on $P[X \geq 11]$.

Example 6.18 Solution

In Table 6.1 the MGF of X is $e^{s\mu + s^2\sigma^2/2}$

$$\phi_X(s) = e^{(11s+s^2)/2}.$$

Thus the Chernoff bound is

$$P[X \geq 11] \leq \min_{s \geq 0} e^{-11s} e^{(11s+s^2)/2} = \min_{s \geq 0} e^{(s^2-11s)/2}.$$

To find the minimizing s, it is sufficient to choose s to minimize $h(s) = s^2 - 11s$. Setting the derivative $dh(s)/ds = 2s - 11 = 0$ yields $s = 5.5$. Applying $s = 5.5$ to the bound yields

$$P[X \geq 11] \leq e^{(s^2-11s)/2}\Big|_{s=5.5} = e^{-(5.5)^2/2} = 2.7 \times 10^{-7}.$$

Based on our model for adult heights, the actual probability (not shown in Table 3.2) is $Q(11 - 5.5) = 1.90 \times 10^{-8}$.

Quiz 6.8

In a subway station, there are exactly enough customers on the platform to fill three trains. The arrival time of the nth train is $X_1 + \cdots + X_n$ where $X_1, X_2, \ldots$ are iid exponential random variables with $E[X_i] = 2$ minutes. Let W equal the time required to serve the waiting customers. For $P[W > 20]$, the probability W is over twenty minutes,

(1) Use the central limit theorem to find an estimate.

(2) Use the Chernoff bound to find an upper bound.

(3) Use Theorem 3.11 for an exact calculation.

Quiz 6.8 Solution

The train interarrival times X_1, X_2, X_3 are iid exponential (λ) random variables. The arrival time of the third train is

$$W = X_1 + X_2 + X_3.$$

In Theorem 6.11, we found that the sum of three iid exponential (λ) random variables is an Erlang $(n = 3, \lambda)$ random variable. From Appendix A, we find that W has expected value and variance

$$E[W] = 3/\lambda = 6 \qquad \text{Var}[W] = 3/\lambda^2 = 12$$

(1) By the Central Limit Theorem,

$$P[W > 20] = P\left[\frac{W-6}{\sqrt{12}} > \frac{20-6}{\sqrt{12}}\right] \approx Q(7/\sqrt{3}) = 2.66 \times 10^{-5}$$

(2) To use the Chernoff bound, we note that the MGF of W is

$$\phi_W(s) = \left(\frac{\lambda}{\lambda - s}\right)^3 = \frac{1}{(1-2s)^3}$$

The Chernoff bound states that

$$P[W > 20] \le \min_{s \ge 0} e^{-20s}\phi_X(s) = \min_{s \ge 0} \frac{e^{-20s}}{(1-2s)^3}$$

To minimize $h(s) = e^{-20s}/(1-2s)^3$, we set the derivative of $h(s)$ to zero: [Continued]

Quiz 6.8 Solution (continued)

$$\frac{dh(s)}{ds} = \frac{-20(1-2s)^3 e^{-20s} + 6e^{-20s}(1-2s)^2}{(1-2s)^6} = 0$$

This implies $20(1-2s) = 6$ or $s = 7/20$. Applying $s = 7/20$ into the Chernoff bound yields

$$P[W > 20] \le \left. \frac{e^{-20s}}{(1-2s)^3} \right|_{s=7/20} = (10/3)^3 e^{-7} = 0.0338$$

(4) Theorem 3.11 says that for any $w > 0$, the CDF of the Erlang $(\lambda, 3)$ random variable W satisfies

$$F_W(w) = 1 - \sum_{k=0}^{2} \frac{(\lambda w)^k e^{-\lambda w}}{k!}$$

Equivalently, for $\lambda = 1/2$ and $w = 20$,

$$P[W > 20] = 1 - F_W(20)$$
$$= e^{-10}\left(1 + \frac{10}{1!} + \frac{10^2}{2!}\right) = 61e^{-10} = 0.0028$$

Although the Chernoff bound is relatively weak in that it overestimates the probability by roughly a factor of 12, it is a valid bound. By contrast, the Central Limit Theorem approximation grossly underestimates the true probability.